November 2008

Solving the Credit Assignment Problem: The Interaction of Explicit and Implicit Learning with Internal and External State Information

Wai-Tat Fu
University of Illinois at Urbana-Champaign

John R. Anderson
Carnegie Mellon University

Published In
Psychological Research, 72, 3, 321-330.

Solving the Credit Assignment Problem: The interaction of Explicit and Implicit learning with Internal and External State Information

Wai-Tat Fu (wfu@cmu.edu)
Human Factors Division and Beckman Institute
University of Illinois at Urbana-Champaign
1 Airport Road, Savoy, IL 61874, USA

John R. Anderson (ja+@cmu.edu)
Department of Psychology
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA

Abstract

In most problem-solving activities, feedback is received at the end of an action sequence. This creates a credit-assignment problem where the learner must associate the feedback with earlier actions, and the interdependencies of actions require the learner to either remember past choices of actions (internal state information) or rely on external cues in the environment (external state information) to select the right actions. We investigated the nature of explicit and implicit learning processes in the credit-assignment problem using a probabilistic sequential choice task with and without external state information. We found that when explicit memory encoding was dominant, subjects were faster to select the better option in their first choices than in the last choices; when implicit reinforcement learning was dominant subjects were faster to select the better option in their last choices than in their first choices. However, implicit reinforcement learning was only successful when distinct external state information was available. The results suggest the nature of learning in credit assignment: an explicit memory encoding process that keeps track of internal state information and a reinforcement-learning process that uses state information to propagate reinforcement backwards to previous choices. However, the implicit reinforcement learning process is effective *only* when the valences can be attributed to the appropriate states in the system – either internally generated states in the cognitive system or externally presented stimuli in the environment.

Introduction

Consider a person navigating in a large office building. The person has to decide when to turn left or right at various hallway intersections. The sequence of decisions is interdependent – e.g., turning left at a particular hallway intersection will affect the decisions at the next intersections. The person may therefore need to keep track of previous actions to inform what actions to take in the future. In reality, memory of previous actions (internal state information) may not be necessary as people can explicitly seek information in the environment (external state information) to know where one is located or which direction to go to reach a destination (Fu & Gray, 2006). Learning to navigate is therefore likely to involve both the retention of internal state information (memory) and the recognition of external state information (signs on the walls). Indeed, many have argued that real-world skills often involve the interplay between cognition (internal), perception, and action (external) that the understanding of these interactive skills requires careful study of how internal (memory) and external information (cues in the environment) are processed in the learning processes (Ballard, 1997; Fu & Gray, 2000; 2004; Gray & Fu, 2004; Larkin, 1989; Gray, Sims, Fu, & Schoelles, in press).

The navigation problem above is an example of one of the most difficult situations in skill learning: when the learner has to perform a sequence of actions but only gets feedback on their success at the end of the sequence (e.g., when the destination is reached). This creates a credit-assignment problem, in which the learner has to assign credits to earlier actions that are responsible for eventual success. When actions are interdependent, either memory of previous actions or recognition of the correct problem state in the external environment is required to properly assign credits to the appropriate actions. In this article, we present results from an experiment in which we study how people learn to solve the credit-assignment problem in a simple but challenging example of such a situation. Our focus is on the recent proposal that humans exhibit two distinct learning processes and we apply it to learning of action sequences with delayed feedback: an explicit process (with awareness) that requires memory for actions and outcomes, and an implicit process (without awareness) that does not require such memory. We will first review research in some related areas that informed the design of our experiment.

Explicit and Implicit Learning

Probability Learning and Classification

There have been numerous studies on the learning of the probabilistic relationship between choices and their consequences. The simplest situation is the probability-learning experiment in which subjects guess which of the alternatives occurs and then receives feedback on their guesses (e.g., Estes, 1964). One robust finding is that subjects often "probability match"; that is, they will choose a particular alternative with the same probability that it is reinforced (e.g., Friedman et al., 1964). This leads many to propose that probability matching is the result of an implicit habit-learning mechanism that accumulates information about the probabilistic structure of the environment (e.g., Graybiel, 1995). One important characteristic of this kind of habit learning is that information is acquired gradually across many trials, and seems to be independent of declarative memory as amnesic patients were found to perform normally in a probabilistic classification task (Knowlton, Squire, Gluck, 1994). However, for non-amnesic human subjects, it is difficult to determine whether probabilistic classification is independent of the use of declarative memory. Since declarative memory is dominant in humans, it has been argued that learners often initially engage in explicit memory encoding in which they seek to remember sequential patterns even when there are none (Yellott, 1969). Researchers argue that true probabilistic trial-by-trial behavior only appears after hundreds of trials – perhaps by then subjects give up the idea of explicitly encoding patterns and the implicit habit-learning process becomes dominant (Estes, 2002; Vulkan, 2000).

Recent research on complex category learning has also provided interesting results suggesting multiple learning systems (Allen and Brooks, 1991; Ashby, Queller, and Berretty, 1999; Waldron and Ashby, 2001). For example, Waldron and Ashby (2001) showed that while a concurrent Stroop task significantly impaired learning of an explicit rule that distinguished between categories by a single dimension, but did not significantly delay learning of an implicit rule that requires integration of information from multiple dimensions.

Sequence Learning

The explicit/implicit distinction has also been investigated through a paradigm called sequence learning (e.g., Cleeremans & McClelland, 1991; Cohen, Ivry, Keele, 1990; Curran & Keele, 1993; Mathews, et al., 1989; Nissen & Bullemer, 1987; Sun, Slusarz, Terry, 2005; Willingham, Nissen, & Bullemer, 1989). In a typical experiment subjects have to press a sequence of keys as indicated by a sequence of lights. A certain pattern of button presses recurs regularly and subjects give evidence of learning this sequence by being able to press the keys for this sequence faster than a random sequence. Although there have been slightly different definitions to capture the details of the implicit/explicit distinction, the key factor seems to be the idea that implicit learning occurs as a facilitation of test performance without concurrent awareness of what is being learned (Reber, 1989; Sun, et al., 2005; Willingham, 1998, but see Shanks & St. John, 1994). However, there seems to be a limit on what the implicit process can learn. For example, Cohen et al. (1990), found that when explicit learning is suppressed by a distractor task, subjects could only learn simple pairwise transitions, but failed to learn higher order hierarchical structures in the sequence.

In neither probability learning nor the typical sequence-learning task is there any doubt about the correctness of a single action. In probability learning there is a single action after which feedback is received. In the typical sequence learning experiment there is a sequence of actions but there is immediate feedback after each action and usually a deterministic relationship between response and correctness. Neither of these paradigms then reflects the complexity of the credit-assignment problem that people frequently face in real life. We combine research from both areas by studying how people learn to assign credits to different actions in a probabilistic sequential choice task, in which sequences of actions are executed before feedback on whether they are correct or not is received, and a particular action sequence is correct only with a certain probability.

Reinforcement Learning

Learning from delayed feedback often involves the temporal credit-assignment problem in which learners must apportion credit and blame to each of the actions that resulted in the final outcome of the sequence. The temporal credit assignment problem is often done by some form of reinforcement learning (e.g., Sutton & Barto, 1998). Recently, psychological research have found that in many learning situations, neural activities in the basal ganglia correlate well with the predictions of reinforcement learning (e.g., Schultz, Dayan, & Montague, 1997). Elsewhere we also show that it produces a wide range of behavioral data in the probability-learning literature and in other delayed feedback learning situations (Fu & Anderson, in press). The role of the basal ganglia is also closely related to the habit-learning (procedural) system in which past response-outcome information is accumulated through experience (e.g., Graybiel, 1995). Such learning is also believed to be distinct from the explicit memory (declarative) system (e.g., Poldrack, et al., 2001; Daw, Niv, & Dayan, 2005).

The basic prediction of reinforcement learning is that when feedback is received after a sequence of actions, only the last action in the sequence will receive feedback but that on later trials its value will then propagate back to early actions. By itself this mechanism cannot learn in cases where success depends on the sequence of actions rather than the individual actions. Memories of previous actions or observations are required to disambiguate the states of the world (e.g., McCallum, 1995). This implies that the cognitive agent needs to explicitly adopt some forms of memory encoding strategies to retain relevant information in memory for future choices.

In our experiment, we study the implicit reinforcement learning process and the explicit memory process in a probabilistic sequential choice task. The task is specifically designed to distinguish between the two processes and we have strong predictions about the outcome in the two condition: When the implicit reinforcement learning process is dominant, learning of items closer to the feedback will be faster than those farther away. When the explicit memory encoding process is dominant, learning of items presented earlier will be faster. We also predict that implicit learning requires distinct state information to propagate credits back to earlier choices. In other words, when state information is absence, implicit learning will fail to learn the dependency between actions.

The Experiment

A probabilistic sequential choice task is designed in which we predict different behavioral patterns when subjects are engaged in explicit and implicit learning processes.

Subjects were told that they were in a room and they had to choose one of the two colors presented on the screen to go to the next room. After making two choices, subjects would either reach an exit or a dead-end. Subjects were instructed to choose the colors that would lead them to the exit as often as possible. Figure 1 shows an example of the task. In room 1, if they chose "red" they would go to room 2 with probability 0.8 and to room 3 with probability 0.2. The probabilities were reversed if "blue" was chosen. After the first choice, if subjects were in room 2, if they choose "yellow" there was a 0.6 probability of going to an exit and 0.4 probability of going to a dead end. Again, the probabilities were reversed if "green" was chosen. If subjects were in room 3, choosing "yellow" would lead to an exit with probability 0.2 and to a dead end with probability 0.8. Choosing "green" would lead to an exit with probability 0.4 and that to a dead end with probability 0.6. Note that if "red" is chosen, "yellow" is more likely to lead to an exit than "green"; but if "blue" is chosen, "green" is more likely than "yellow". The choice of colors in the second choice is therefore dependent on the first choice.

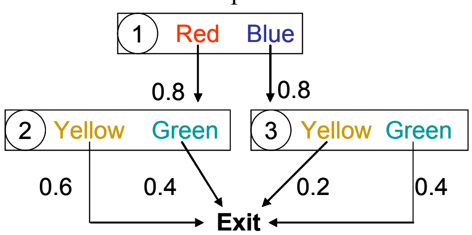

Figure 1. The probabilistic sequential task. The circled numbers represent room numbers, and the numbers next to the arrows represent transition probabilities. Note that in room 3, regardless of what is chosen, there is a higher probability that it will lead to a dead-end compared to room 2. The actual colors were randomly selected from eight colors (red, green, yellow, blue, brown, gray, magenta, and orange) for each subject.

One strategy in this task was to conduct a "tree-searching" by explicitly encoding the choices in memory and observing their outcomes. In this task, the probabilities were chosen such that, even if subjects randomly chose a color in the second choice, the probability that choosing "red" would eventually lead to an exit was higher than that for choosing "blue" (it can be easily shown that the marginal probabilities were 0.46 and 0.34 for choosing "red" and "blue" respectively). On the other hand, if subjects randomly chose a color in the first choice, the probabilities that choosing "yellow" or "green" would lead to an exit were equal (it can be shown that the marginal probability would both be 0.4). The task was designed such that when engaged in explicit memory encoding and searching, the first choices would be learned faster than the second choice, as it was more likely that the memory traces of the better first choice would be strengthened faster than those for the better second choice.

To study the nature of the implicit learning process, we introduced a "2-back" secondary task to suppress the otherwise dominant explicit memory encoding process. The secondary task required subjects to listen to a continuous stream of numbers (from 0 to 9) from the speakers. Starting from the third number, subjects had to press the control key on the keyboard if the number is identical to the numbers two numbers before. For example, if they heard the numbers 0, 3, 2, 3, and 0, they had to press the control key the second time they heard 3. The numbers were presented once every two seconds. Subjects had to maintain their performance at 80% or better at the 2-back task while performing the probabilistic sequential task.

From earlier discussion, the basic prediction of the implicit reinforcement-learning process is that actions close to the feedback will acquire value first and then their value will propagate back to early actions. Thus, in contrast to the explicit memory encoding process, learning of the choices closer to the feedback will be faster than earlier choices. However, in the probabilistic sequential choice task, since the choices were designed to be dependent, it was impossible to learn the second choice before learning which color was better in first choice. We therefore need to provide some external state information for subjects to learn to recognize their current state in the second choice (i.e., whether they were in room 2 or room 3 in Figure 1), so that it is possible for them to learn the second choice before the first choice as predicted by the implicit reinforcement learning process. In addition, since the implicit learning process does not require explicit memory encoding, the prediction is that subjects may be able to learn to choose the more likely colors without concurrent awareness of them.

To study the effect of external state information on the learning of the two choices, we placed half of the subjects in the *distinct* condition and the other half to the *ambiguous* condition. In the distinct condition, in addition to the two colors, there was also a distinct object in room 2 and 3 (e.g., a computer in room 2 and a telephone in room 3). Subjects did not see the object in the ambiguous condition. Our expectation was that in the distinct condition, the presence of the object would help subjects to identify which room they were in. This would allow them to choose the more likely color in the second choice set even without explicit memory of their first choice. In the ambiguous condition, choosing the more likely second color would require internal state information encoded by explicit memory.

Method

52 subjects in the Carnegie Mellon University community were recruited for the experiment. Four of the subjects could not maintain the 2-back task performance at 80% and were excluded. Subjects received a base payment of $8 plus a bonus payment of up to $7 depending on performance. Half of the remaining 48 subjects were assigned to the single-task group and the other half to the dual-task group; and subjects in each group were further divided into the distinct and ambiguous conditions. Subjects started with an initial score of 10 points. When an exit was reached, 5 points would be added to the final score; when a dead-end was reached, 1 point would be deducted from the final score. Subjects were paid one cent for each point in the total score for the bonus payment. Each subject finished 20 10-trial blocks. At the end of the experiment, subjects were asked to write down any strategy they used and whether they thought that any of the colors was more likely lead to the exit.

Results

Subjects who could write down the more likely colors in all three rooms (thus the choice dependency) were placed in the aware group; otherwise they were placed in the not-aware group (see Table 1). In the dual task condition, most of the subjects could not write down the more likely colors in any of the rooms, while subjects in the single task condition could write down the more likely colors in at least two of the rooms (we chose not to include them in the aware group as they apparently were not aware of the choice dependency between the two choices).

Table 1. Number of subjects who wrote down the more likely colors in each of the experimental condition. All = all rooms, none = none of the rooms, R1 = room 1 only, and R1 & R2 = room 1 and 2 only, etc. In the ambiguous condition, subjects were not aware of the distinction of room 2 and 3.

	Single		Dual	
Rooms	Distinct	Ambiguous	Distinct	Ambiguous
All	9	7	2	1
R1 & R2	2	4	0	1
R1 & R3	0	1	0	0
R2 & R3	1	0	2	1
R1	0	0	0	0
R2	0	--	1	--
R3	0	--	0	--
none	0	0	7	9

A 2 (first/second choice) x 2 (awareness) x 2 (single/dual task) x 2 (distinct/ambiguous condition) ANOVA on the choice proportions on the more likely colors shows that the main effects of awareness and condition were significant ($F(1,40)=12.21$, $MSE=0.19$, $p<0.001$; $F(1,40)=5.33$, $MSE=0.19$, $p<0.05$ respectively); learning was better in the aware group than the not-aware group, and was better in the distinct condition than the ambiguous condition. There were significant choice x awareness x condition and choice x awareness interactions ($F(1,40)=8.79$, $MSE=0.088$, $p < 0.01$ and $F(1,40)=18.68$, $MSE=0.088$, $p < 0.001$ respectively). No other interaction involving choice was significant. Since the main effect of task was not significant ($F(1,40)=0.95$, $MSE=0.21$, $p=0.34$), nor was any of its interaction, the results were collapsed across tasks in Figure 2, which shows the mean choice proportions of the more likely colors in each 20-trial block. Consistent with our expectation, in the distinct condition, subjects in the aware group learned the first choice faster than the second choice while subjects in the not-aware group learned the second choice faster than the first choice. In the ambiguous condition, subjects in the aware group also learned the first choice faster than the second choice. However, in contrast to the distinct condition, subjects in the not-aware group were not significantly above chance throughout the 10 20-trial blocks for, indicating that they failed to learn implicitly when state information was absent.

The main effect of blocks was significant ($F(9,360)=6.86$, $MSE= 0.019$, $p < 0.001$). The blocks x awareness x condition interaction was significant ($F(9,360)=3.70$, $MSE=0.019$, $p < 0.001$). No other interaction involving blocks was significant. The significant interaction could be explained by the fact that, except the not-aware group in the ambiguous condition, subjects significantly increased their choice proportions of the more likely colors across trials. Indeed, the last four blocks of both choices were significantly above chance for all but the not-aware group in the ambiguous condition.

The results were consistent with the proposed distinct learning processes in the probabilistic sequential choice task. As reflected by our awareness measure, in the single-task condition, most of the subjects explicitly remembered the outcomes of the choices and were aware of the choice dependencies. Consistent with our expectation, subjects in the aware group presumably conducted a tree-searching strategy, and learned the first choice faster than the second choice. In the dual-task condition, since the explicit encoding of past experiences was suppressed, most of the subjects were not aware of the most likely colors. Nevertheless, in the distinct condition, subjects increasingly selected the more likely colors, demonstrating learning of the dependency between the choices[1]. Consistent with the reinforcement-learning mechanism, learning of the second choice was faster than the first choice, despite the asymmetry of choice probabilities in the design of the task. The result also suggests that reinforcement learning does not require explicit memory encoding and concurrent awareness to learn the choice dependency.

In the ambiguous condition, the dependency between choices could only be learned if subjects remembered the first choice when making the second choice. Most subjects in the single-task condition were aware of the better colors in both choices and chose them increasingly often across trials. This suggests that subjects in the aware group did learn the dependency of choices. Similar to the subjects in the aware group in the distinct condition, learning of the first choice was faster than the second choice. In the dual-task condition, the suppression of the memory encoding of the first choice significantly hampered the discovery of the dependency. Subjects failed to learn to choose the better colors above chance level. Apparently, reinforcement learning failed when the final states (i.e., room 2 and room 3) were indistinguishable, as both internal and external cues were not available. It suggests distinct states information is essential for the proper propagation of credits to earlier state-action pairs.

.

[1] Note that if subjects were not aware of the choice dependency and always chose one of the more likely colors in the second choice set (i.e., chose “yellow” in both room 2 and 3 using the example shown in Figure 1), the choice proportion would have been 80% of the choice proportion of the more likely color in the first choice (i.e., approximately 0.8 x 0.8 = 0.64 in the last 3 blocks). Since the second choice proportions were higher than 0.64, subjects had learned to choose the more likely colors in both room 2 and room 3 – i.e., they had learned the dependency between the choices.

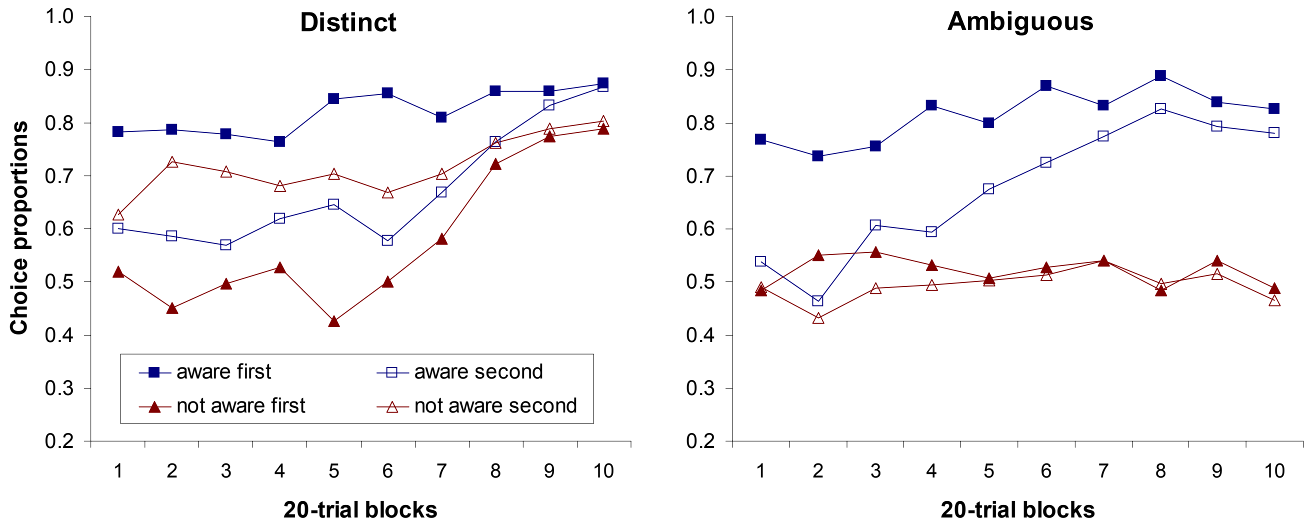

Figure 2. Choice proportions of the colors that were more likely to lead to the exit in the distinct and ambiguous conditions in each of the 20-trial blocks. Using the example shown in Figure 1, "first" would be the choice proportions of "red", and "second" would be the sum of the choice proportions of "yellow" and "green" in room 2 and room 3 respectively.

Discussions

The primary questions addressed by the study are (1) whether there are explicit and implicit modes of learning in probabilistic sequential choice tasks, as suggested by the literature on probability learning and sequence learning, if so (2) whether the implicit learning process is consistent with the credit-assignment mechanism in reinforcement learning, and (3) whether explicit external state information is required to propagate credits back to earlier actions when the actions are interdependent as predicted by the reinforcement learning process. Results from the experiment seem to answer all three questions in the affirmative.

In an uncertain environment, people learn to choose the right actions by identifying states of the cognitive system and the environment associated with positive and negative valence. In most situations, the states consist of combinations of internally encoded responses and externally presented stimuli. In most situations, the explicit, goal-directed tree-searching strategy seems dominant, which allows people to encode responses and their outcomes internally. The internally encoded state information then guides future selection of actions. We found that in addition to this dominant explicit encoding process, an implicit reinforcement learning process allows learning by monitoring the outcomes of responses (positive or negative valences). However, this implicit reinforcement learning process is effective *only* when the valences can be attributed to the appropriate states in the system – either internally generated states in the cognitive system or externally presented stimuli in the environment.

The probabilistic sequential choice task used in the experiments, although simple, contains essential components in interactive skill learning, in which a sequence of actions are performed before reinforcement on the full course of action is received. Solving the credit-assignment problem is crucial for learning in this kind of situation, as the delayed feedback has to propagate back to the appropriate actions that are responsible for the desirable or undesirable outcome. The reinforcement-learning process provides a straightforward explanation of how feedback propagates back to earlier actions. Initially, only the action that leads to outcome gets credit or blame. The next time some of that credit/blame propagates back to the previous actions. Eventually, credit/blame can find its way back to critical early actions in a long chain of actions leading to a reward. The effectiveness of this process, however, depends on whether the effects of these actions are independent of each other. When the actions are interdependent, either distinct external state information or memory of earlier actions is required to ensure the proper assignment of credits for effective skill learning.

Acknowledgment

The current work is supported by a grant from the Office of Naval Research (N00014-99-1-0097).

References

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics, 61*, 1178-1199.

Allen, S.W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120*, 3-19.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(4), 723-742.

Cleeremans, A. & McClelland, J.L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology : General*, 120, 235-253.

Cohen, A., Ivry, R., & Keele, S. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189–202.

Daw, N., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*, 1704-1711.

Estes, W. K. (1964). Probability learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.

Estes, W. K. (2002). Traps in the route to models of memory and decision. *Psychonomic Bulletin and Review, 9* (1), 3-25.

Friedman, M. P., Burke, C. J., Cole, M., Keller, L., Millward, R. B., & Estes, W. K. (1964). Two-choice behavior under extended training with shifting probabilities of reinforcement. In R. C. Atkinson (Ed.), *Studies in mathematical psychology* (pp. 250-316). Stanford, CA: Stanford University Press.

Fu, W. & Anderson, J. (in press). From recurrent choice to skill learning: A model of reinforcement learning. *Journal of Experimental Psychology: General.*

Fu, W. & Gray, W. D. (2000). Memory versus Perceptual-Motor Tradeoffs in a Blocks World Task. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Erlbaum.

Fu, W. & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science, 28* (6).

Fu, W. & Gray, W. D. (2006). Suboptimal Tradeoffs in Information-Seeking. *Cognitive Psychology*. *52* (3), 195-242.

Gray, W. D., & Fu, W. (2004). Soft Constraints in Interactive Behavior: The Case of Ignoring Perfect Knowledge In-The-World for Imperfect Knowledge In-The-Head .*Cognitive Science, 28* (3), 359-382.

Gray, W. D., & Sims, C., Fu, W., Schoelles, M. (in press). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review.*

Graybiel, A.M. (1995). Building action repertoires: memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology, 5*, 733-741.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399-1402.

Knowlton, B. J., Squire, L. R., & Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1*, 106-120.

Larkin, J. H. (1989). Display-based problem solving. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 319–341). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mathews, R., Buss, R., Stanley, W., Blanchard-Fields, F., Cho, J., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1083–1100.

McCallum, A. K. (1995) *Reinforcement Learning with Selective Perception and Hidden State*, PhD. Thesis, University of Rochester.

Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.

Poldrack, R., Clark, J., Pare-Blagoev, E., Shohamy, D., Moyano, J., Myers, C., & Gluck, M. (2001). Interactive memory systems in the human brain. *Nature, 414*, 546-550.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General 118*, 219-235.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593-1599.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367-447.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review, 112*, 159-192.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Vulkan N. 2000. An economist's perspective on probability matching. *Journal of Economic Surveys, 14*, 101–118

Waldron, E., & Ashby, G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. Psychonomic Bulletin & Review, 8, 168-176.

Willingham, D. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*, 558-584.

Willingham, D., Nissen, M., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1047-1060.

Yellott, J. L. (1969). Probability learning with noncontingent success. *Journal of mathematical psychology, 6*, 541-575